DECISION-MAKING IN THE
FACE OF UNCERTAINTY – II

Richard Bellman

P-593

8 November 1954

The RAND Corporation
1700 MAIN ST · SANTA MONICA · CALIFORNIA

SUMMARY

We consider multi-stage processes involving both zero-sum and non-zero sum games. Using the concept of "games of survival" we derive approximate solutions for both classes of multi-stage games under various realistic assumptions.

DECISION-MAKING IN THE FACE OF UNCERTAINTY - II

by

Richard Bellman

§1. Introduction

In the first paper of this series, [1], we considered a multi-stage decision process of simple and yet general type and showed that an approximate solution of plausible and intuitive sort could be obtained under certain reasonable assumptions.

In this paper, we shall consider a more difficult class of problems, involving conflict between two groups. In many situations of importance, one group may be considered to be the inanimate forces of nature.*

* Summarized by the phrase "the perversity of inanimate objects".

When the two groups are in direct conflict we have the theory of zero-sum games to assist us. When the two groups are partly apposed and partly not, we encounter the problems of non-zero sum games, where there is essentially no theory to guide us. Nevertheless, we shall show that under certain conditions we can once again obtain approximate modes of procedure which seem eminently reasonable.

§2. Description of the Multi-Stage Process

We shall consider a multi-stage process of the following type, which we shall call a "game". There are two protagonists, whom we shall call "players", named rather prosaically P and Q. The first of these players, P, has a choice of M different plays, which we shall designate by the numbers 1,2,...,M, and the second player, Q, has a choice of N different plays, denoted by 1,2,...,N. If P chooses the i^{th} of his choices and Q the j^{th} of his choices, then P receives a quantity a_{ij} and Q a quantity b_{ij}. These quantities may be negative, in which case a loss is sustained.

We can now consider this situation to repeat itself for a finite or unbounded number of stages and pose the problem of determining the optimal mode of play for each player under the assumption that each wishes to maximize his over-all return.

§3. Zero-Sum Games - Single Stage

The simplest case, and unfortunately the only case in which a satisfactory theory exists, is that where $b_{ij} = -a_{ij}$, which is

to say the players are in direct opposition since one's gain is the other's loss.

We shall begin with the discussion of the single-stage process. It is clear that the determination of optimal play is trivial if either of the players is required to move before the other. The only interesting case is that where both players are required to move simultaneously.

In these circumstances they protect themselves by mixing their choices. Les us assume then that P chooses the $i^{\underline{th}}$ play with probability x_i and Q his $j^{\underline{th}}$ choice with probability y_j. The vector $x = (x_1, x_2, \ldots, x_n)$ specifies P's probability distribution, and $y = (y_1, y_2, \ldots, y_n)$ specifies Q's probability distribution.

The expected return for P will be

$$(1) \quad E_P(x,y) = \sum_{i,j=1}^{N} a_{ij} x_i y_j \quad ,$$

and the expected return for Q, $E_Q(x,y)$ will be the negative of this,

$$(2) \quad E_Q(x,y) = -E_P(x,y)$$

The first player, P, will choose his probability distribution x so as to maximize E_P and the second player Q will choose y so as to minimize E_P.

We can then define two values

$$(3) \quad V_P = \underset{y}{\text{Min}} \; \underset{x}{\text{Max}} \; E_P(x,y)$$

and

(4) $V_Q = \underset{x}{\text{Max}}\ \underset{y}{\text{Min}}\ E_P(x,y)$

In each case the variation is over the regions defined by

(5) (a) $x_i \geq 0, \quad \sum_{i=1}^{M} x_i = 1,$

(b) $y_j \geq 0, \quad \sum_{j=1}^{N} y_j = 1.$

The first value, V_P, is the expected return to P if Q is required to choose y before P chooses x, and the second is the expected return to P if P must choose x before Q chooses y.

It is a remarkable fact, the celebrated min-max theorem of Von Neumann, [4], the basic theorem of the theory of games, that

(6) $V_P = V_Q.$

The interpretation of this result is that P can announce x in advance, and Q likewise, without either gaining from this advance knowledge.

§4. Zero-Sum Games – Finite Resources

In many situations, involving multi-stage play, the above model is not satisfactory. This is particularly true in multi-stage processes where both sides have finite resources. Here the game automatically terminates when either player has no resources.

Let us assume then that each side now plays to ruin the other,

with the game continuing until one or the other player is bankrupt. Let p represent the initial amount possessed by P and q the initial amount possessed by Q. We define

(1) $f(p,q)$ = the probability that P survives Q when P starts with p, Q with q and both sides use optimal play.

Games of this variety are aptly called "games of survival", cf [2], [3].

§5. Games of Survival - Mathematical Formulation

Since P wins what Q loses, and vice versa, the total quantity of resources in the game remains constant, and equal to $p_0 + q_0 = N$, the initial total, to specify the state of the game it is sufficient then to state the amount of resources possessed by P.

We replace $f(p,q)$ by the function of one variable $f(p)$. Let us now derive a functional equation for $f(p)$. Enumerating the possible outcomes of one stage of play, we see that

$$(1) \quad f(p) = \sum_{i,j=1}^{N} x_i y_j f(p+a_{ij}) \quad ,$$

where $x = (x_1, x_2, \ldots, x_M)$ and $y = (y_1, y_2, \ldots, y_N)$ are the probability distributions of P and Q determining the initial play.

Since P plays to maximize $f(p)$ and Q to minimize $f(p)$, we obtain the equation

$$(2) \quad f(p) = \operatorname*{Max}_x \operatorname*{Min}_y \sum_{i,j=1}^{N} x_i y_j f(p+a_{ij})$$

$$= \operatorname*{Min}_y \operatorname*{Max}_x \sum_{i,j=1}^{N} x_i y_j f(p+a_{ij}) \quad .$$

This equation is valid for $0<p<N$. We also have

(3) $f(p) = 1, \qquad p \geq N$

$f(p) = 0, \qquad p \leq 0.$

§6. Approximate Solution

A solution of (2) yields $f(p)$ and thus the vector $x(p)$, $y(p)$, the probability distributions determining the initial play. We wish to determine approximate values for $x(p)$ and $y(p)$ under certain assumptions concerning the sequence of pay-offs, (a_{ij}).

We assume that no single play creates any appreciable change in the state of the game, which is to say that a_{ij} is small compared to p.

We then write

(1) $f(p+a_{ij}) = f(p) + a_{ij}f'(p)$.

Then (5.2) takes the form

(2) $$f(p) \cong \underset{x}{\text{Max}}\ \underset{y}{\text{Min}} \left[\sum_{i,j=1}^{M,N} x_i y_j (f(p) + a_{ij}f'(p))\right]$$
$$\cong \underset{y}{\text{Min}}\ \underset{x}{\text{Max}} \left[\sum_{i,j=1}^{M,N} x_i y_j (f(p) + a_{ij}f'(p))\right],$$

or

(3) $$0 \cong \underset{x}{\text{Max}}\ \underset{y}{\text{Min}} \left[f'(p) \sum_{i,j=1}^{M,N} a_{ij}x_i y_j\right]$$
$$\cong \underset{y}{\text{Min}}\ \underset{x}{\text{Max}} \left[f'(p) \sum_{i,j=1}^{M,N} a_{ij}x_i y_j\right].$$

Let us now assume further that, as is true in all realistic situations, it definitely pays P to have a larger initial quantity of resources, that is to say $f'(p) > 0$ for all p.

Then (3) is replaced by

$$(4) \quad c \cong \underset{x}{\mathrm{Max}} \ \underset{y}{\mathrm{Min}} \left[\sum_{i,j=1}^{M,N} a_{ij} x_i y_j \right]$$

$$\cong \underset{y}{\mathrm{Min}} \ \underset{x}{\mathrm{Max}} \left[\sum_{i,j=1}^{M,N} a_{ij} x_i y_j \right] .$$

The meaning of this equation is that for large p, with a large number of plays remaining until the end of the game, the play is approximately the same as that employed in the single-stage process where both players wish merely to maximize the expected return from one play.

This approximate solution has precisely the same structure as that given for the one-person process in [1].

The important feature of (4) is that we obtain the same approximate equation regardless of the significance of f(p). Consequently, even in situations where f(p) is not completely determined, as frequently occurs in realistic situations, we know that we possess a good approximation to optimal play.

§7. Non-Zero Sum Games - Games of Survival

Let us now turn to a discussion of the more general situation where $b_{ij} \neq -a_{ij}$. Here there is no theory for the determination of optimal play in a single-stage process. Consequently, we shall

turn immediately to the discussion of multi-stage processes. We assume once more that both players strive to ruin the other and continue the game until this occurs.

Let p be the initial amount possessed by P and q the amount possessed by Q, and introduce the function defined by (4.1).

This function satisfies the functional equation

$$(1)\quad (a)\quad f(p,q) = \underset{x}{\text{Max}}\ \underset{y}{\text{Min}} \sum_{i,j=1}^{M,N} x_i y_j f(p + a_{ij}, q + b_{ij})$$

$$= \underset{y}{\text{Min}}\ \underset{x}{\text{Max}} \sum_{i,j=1}^{M,N} x_i y_j f(p + a_{ij}, q + b_{ij}), \quad p,q>0,$$

$$(b)\quad f(p,q) = 1, \quad p > 0, \quad q \leq 0$$

$$= 0, \quad p \leq 0, \quad q > 0$$

$$= 1/2, \quad p = q = 0 .$$

§8. Approximate Solution

Let us assume that a_{ij} and b_{ij} are both negative, so that we are dealing with an attrition process, and that a_{ij} and b_{ij} are small compared to p and q.

We write

$$(1)\quad f(p+a_{ij}, q+b_{ij}) = f(p,q) + a_{ij}f_p + b_{ij}f_q \quad .$$

Then (7.1a) yields

$$(2)\quad 0 \cong \underset{x}{\text{Max}}\ \underset{y}{\text{Min}} \left[f_p \sum_{i,j=1}^{N} a_{ij}x_i y_j + f_q \sum_{i,j=1}^{M,N} b_{ij}x_i y_j \right]$$

$$\cong \underset{y}{\text{Min}}\ \underset{x}{\text{Max}} \left[f_p \sum_{i,j=1}^{N} a_{ij}x_i y_j + f_q \sum_{i,j=1}^{N} b_{ij}x_i y_j \right]$$

As before we assume that $f_p > 0$, $f_q < 0$. Since $\Sigma\, a_{ij}x_iy_j$, $\Sigma\, b_{ij}x_iy_j$ are both negative, (2) yields

$$(3)\quad \frac{f_p}{f_q} \cong \underset{x}{\text{Max}}\ \underset{y}{\text{Min}} \left[\sum_{i,j=1}^{M,N} b_{ij}x_iy_j \Big/ \sum_{i,j=1}^{M,N} a_{ij}x_iy_j\right]$$

$$\cong \underset{y}{\text{Min}}\ \underset{x}{\text{Max}} \left[\sum_{i,j=1}^{M,N} b_{ij}x_iy_j \Big/ \sum_{i,j=1}^{M,N} a_{ij}x_iy_j\right]$$

The interpretation of this equation is that both sides play approximately so as to maximize or minimize respectively the ratio

$$(4)\quad R(x,y) = \sum_{i,j=1}^{M,N} b_{ij}x_iy_j \Big/ \sum_{i,j=1}^{M,N} a_{ij}x_iy_j$$

That, under the assumptions on a_{ij} and b_{ij},

$$(5)\quad \underset{x}{\text{Max}}\ \underset{y}{\text{Min}}\ R(x,y) = \underset{y}{\text{Min}}\ \underset{x}{\text{Max}}\ R(x,y)$$

is a theorem also due to Von Neumann, and recently established in a different manner by Shapley, [5].

§9. A Rationale for Non-Zero Sum Games

The importance of the above result resides in the fact that it furnishes us a motive for using $R(x,y)$ as a universal pay-off function for non-zero sum games. Whether or not it is to be accepted in any particular situation will depend on other properties of the game.

BIBLIOGRAPHY

1. Bellman, R., "Decision-Making in the Face of Uncertainty - I", Navy Quarterly of Logistics (to appear).

2. ———————, "An Introduction to the Theory of Dynamic Programming", RAND Report No. R-245, 1953.

3. Bellman, R. and LaSalle, J., "Non-Zero Sum Games and Stochastic Processes".

4. Von Neumann, J. and Morgenstern, O., "The Theory of Games and Economic Behavior", Princeton University Press, 1948.

5. Shapley, L., "Stochastic Games", Proc. Nat. Acad. Sci., Vol. 39 (1953), pp. 1095-1100.

/md